CIS 3362  
8/31/22

substitution ➞ breaker
null char, code words ➞ adjustment

Vigenere Cipher (Unbreakable Code!)

Plaintext: IT IS VERY HOT OUTSIDE
Keyword: KNIGHTSTKNIGHTSKNIG

8 19 8 18 21 4 17 24
+ 10 13 8 6 7 19 18 10

18 32 16 24 28 23 35 34
6 2 9 8

18, 6, 16, 24, 2, 23, 9, 8
SG Q Y C X J I ...

advantages: (1) easier to remember
*(2) freq info distorted
2 Ts can encrypt to 2 diff letters
2 diff letters can encrypt to same letter

disadvantages: (1) easier to "guess" at a keyword.

TO DO: Code up Java version encryption
Breaking Vigenere

Letter freq not preserved  \[ \text{BUT} \]
Some plaintext chars are shifted by the same amount!

\[ \text{KNIGHTS KNIGHTS KNIGHTS KNIGHTS} \]

Letters pos 0, 7, 14, ...  \[ \text{Same shift} \]
Letters pos 1, 8, 15, ...  \[ \text{Same shift} \]
Letters pos 6, 13, 20, ...  \[ \text{Same shift} \]

Issues - bins will have fewer chars than for reg sub.

- won't get digraph, trigraph into!
How to deal w/ issues

(1) repeated digraphs/trigrams

Imagine I use the same word 5 times.

keyword length was 8.

List 5 random #s from set {0,1,2,3,4,5,6,7} each chosen independently. What's probability they are all different?

\[
6, \underline{3}, 4, \underline{1}, 7
\]

\[
1 \times \frac{7}{8} \times \frac{6}{8} \times \frac{5}{8} \times \frac{4}{8} \approx 0.0206
\]

80% chance that the repeated word is encrypted identically!??

Kasiski - repeated dig/trigram

⇒ likely means same plaintext

⇒ difference in the indexes is divisible by the keyword length.

(1) find all repeats

(2) record all differences in index

(3) GCD all #s in step